

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 13:05:41 ; Search time 45 Seconds
(without alignments)
3814.161 Million cell updates/sec

Title: US-10-027-000-2

Perfect score: 4391

Sequence: 1 MADIDVEALKLKITLAEKVD.....DCVALRCKFTVGETWWMSGV 833

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21;*

1: sp_archea;*

2: sp_bacteria;*

3: sp_fungi;*

4: sp_human;*

5: sp_invertebrate;*

6: sp_mammal;*

7: sp_minc;*

8: sp_organelle;*

9: sp_phage;*

10: sp_plant;*

11: sp_reptile;*

12: sp_virus;*

13: sp_vertebrate;*

14: sp_unclassified;*

15: sp_rvirus;*

16: sp_bacteriaph;*

17: sp_archeap;*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1996.5	45.5	832	3 Q9P6J6
2	1970.3	38.8	862	3 Q9C3Z9
3	1595.5	36.3	861	3 Q9HGK2
4	1143.5	26.0	772	2 006050
5	1095.5	24.9	754	2 008331
6	982	22.4	720	2 060038
7	959	21.8	721	16 Q9WXN2
8	958	21.8	828	2 051843
9	957	21.8	859	16 087852
10	954	21.7	808	2 070021
11	946.5	21.6	762	16 Q9A7G7
12	944.5	21.5	809	2 Q9ZGH5
13	941	21.4	781	2 Q9F3Y0
14	939.5	21.4	793	16 Q8ZCZ9
15	919.5	20.9	732	2 Q9RH03
16	874.5	19.9	752	2 P96316

ALIGNMENTS

RESULT 1

Q9P6J6 PRELIMINARY, PRM: 832 AA.

ID Q9P6J6; AC Q9P6J6; DT 01-OCT-2000 (TREMBrel, 15, Created); DT 01-OCT-2000 (TREMBrel, 15, Last sequence update); DT 01-MAR-2002 (TREMBrel, 20, Last annotation update); DE Beta-D-glucosidase (EC 3.2.1.21), glycosyl hydrolase family 3. SPB01683_04. GN Schizosaccharomyces pombe (Fission yeast). OS Schizosaccharomyces pombe (Fission yeast). OC Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomyces. OC NCBI_TaxID=4896; OX [1] RN RA MCombie W.R. RT "Sequence analysis of a region of the fission yeast genome." RL Submitted (APR-2000) to the EMBL/Genbank/DDBJ databases. DR EMBL; AL335920; CAR91166.1; DR InterPro; IPR02772; GL3C. DR InterPro; IPR01164; GH_3C. DR Pfam; PF00933; Glyco_hydro_3; 1. DR Pfam; PF01915; Glyco_hydro_3_C; 1. DR PRINS; PR00133; GLYDYLASE3. DR PROSITE; PS00775; GLYCOSYL_HYDROL_F3; UNKNOWN_1. KW Hydrolase. SEQUENCE: 832 AA: 93169 MW: 2448ER2B248BB02F CRC64: Query Match Best Local Similarity 45.5%; Score 1996.5; DB 3; Length 832; Matches 389; Conservative 157; Mismatches 282; Indels 9; Gaps 7; QY 1 MADIDVEALKLKITLAEKVDLLAGIDFWHKKALPKHGPPSLRFTDGNGVAGTKFENGIVP 60 Db 1 MMEHDERLNLQDIDSERKAMLSGTDWTAIPRLNIPMSIRLSQDGPNNGTSEPNSSP 60 QY 61 AACFPCTGTSLGSTENQTLIEAGKMMKAEAKSAHVLGPTINMQRSPUGRGFBESIGE 120

RESULT 3	
Q9HGX2	PRELIMINARY; PRT; 861 AA.
Q9HGX2	PRT; 861 AA.
AC	01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	Purutase beta glucosidase (EC 3.2.1.21).
BG1	Agaricus bisporus ('Common mushroom').
GN	Eukaryota: Fungi: Basidiomycota: Hymenomycetes: Agaricales; Agaricaceae; Agaricus. NCBI-TaxID=5341;
OS	{1} SEQUENCE FROM N.A.
RC	STRAIN=Ds49;
RA	Morales-Almora P.; Thurston C.F.;
RT	"Molecular analysis of the cellulolytic genes in Agaricus bisporus.";
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AJ293760; CAC03462; 1; -.
DR	InterPro; IPR002772; GH_3C.
DR	InterPro; IPR01764; GH_3N.
DR	Pfam; PF00933; Glyco_hydro_3_1.
DR	Pfam; PF01915; Glyco_hydro_3_C; 1.
DR	PRINS; PRO0133; GLYDRASE3.
SEQUENCE	861 AA; 94535 MW; 8B52C0A18DB39AB3 CRC64;
Query Match	Best Local Similarity 42.0%; Score 1595.5; DB 3; Length 861; Matches 365; Conservative 147; Mismatches 306; Indels 51; Gaps 22; GQ
Q9HGX2	2 ADIVDEAILEKILKLTAEKVLDLAGIDFWHKKPALKHGVPSLRFDTGPGNGVRGTKEFNGVPA 61
Db	7 ANADDIKTIVDALTDEATIILSAGVGFWHHTAERLQIPAVKVSQDGPNQGRGNIEFMGIFPA 66
QY	62 ACFPGCTSLGSTFNOTLLEEGAG-KMMGKFAIAKSAAHVTGIFTINMORSPLGGFESTGE 120
Db	67 KCLPSSTAMGATEFDRDILLEBVGKLKLLAREAKLRSASVILAPTCNIDORNPLGGSFESIE 126
QY	121 DPFLLAGLGAALIRGIOSTGQVATIKHFLCNDQEDRRMMVOSIVTERALEREIALPFOIA 180
Db	127 DPFLSGMICSSAYINGIQKGIGATIKHFWGNDKDRDGTGDSI-HYRTSGDLSLPSVHA 185
QY	181 VRDSQP- -GFMAYTINGINVCSCBNPKYDGLRKEWGNDKDRDGTGDSI-HYRTSGDLSLPSVHA 238
Db	186 GSRNNPLHGAJMTAYRNVNGIHWENPELQLKILDEWNSDVMIMSDWFGTYSDVGLNA 245
QY	239 GLDLEMPGPFRFRG-ETLKFNNSNGKPFTHIDORAREVLFQVKCAASGVT-E---NGP 293
Db	246 GLDLEMPGKMKWRSLESVNRNSIQSRSKVTARKVKERAKVVELVKKC-AQGAPETLDGDL 304
QY	294 ETTVNTPEAAALLRKVGKEGIVLKNENNVLPLSK - -KTLVGPNAKQAYHGGSA 351
Db	305 ERTL-DSDEALNMRKAFAASIVLVLKNDKVKPLDRKSLSKVATVGGNAKAVLGGSA 363
QY	352 ALRAYAVTFPDGISKQL-ETPP ---SYTVGAVT-VPFLGQCLTPDGAPOM- -RWR 403
Db	364 ALKSYFVVARPAGLVKALKEVNPNEVYSEGARAKTLPDLDLFTESQRGWTGSMH 423
QY	404 VFNEP- - - - -PGTPNRQHDELFPTKTDMHLDVYHYPKADTWADMEGTYADE-DCT 456
Db	424 AHENDDSLVALPPTKQYDE ---TRMFISSSPAGTITKKWLRMKGYLPRERDCK 478
QY	457 YELGIWVCGAKAWYDQVWNTAQVPGDAFGSATRERVTGRINLYKGNTYAFKIEG 516
Db	479 FEGFLTAGKAKLFLVKGKLYDNWTRQRQVEFGCGSIEETGVKAGVKHELYVDFC 538
QY	517 S- - - - -APITLQKDTITPGHGSIRVGCKVQDQAELEKSVLAKEHDQVTCAGLNADWIT 574
QY	{1} SEQUENCE FROM N.A.
RC	STRAIN=Ds49;
RA	Morales-Almora P.; Thurston C.F.;
RT	"Molecular analysis of the cellulolytic genes in Agaricus bisporus.";
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AJ293760; CAC03462; 1; -.
DR	InterPro; IPR002772; GH_3C.
DR	InterPro; IPR01764; GH_3N.
DR	Pfam; PF00933; Glyco_hydro_3_1.
DR	PRINS; PRO0133; GLYDRASE3.
SEQUENCE	772 AA; 86092 MW; 0E6CD88BF7923CC3 CRC64;
Query Match	Best Local Similarity 31.6%; Score 1143.5; DB 2; Length 772; Matches 261; Conservative 134; Mismatches 243; Indels 187; Gaps 14; GQ
Q9HGX2	4 IDVERAILKKLTAEKVLDLAGIDFWHKKPALKHGVPSLRFDTGPGNGVR---GTFKFG 58
Db	13 MDIAHMEINTLLEKASLSGSDAFTWHTKAIERDLPQIWSDGPRLKRNVDSSNDEPEA 72
QY	59 VPAACGFCPGTSLGSTFNOTLLEEGAGKMMGKFAIAKSAAHVTGIFTINMORSPLGGFESTI 118
Db	73 IEAVCFPTAAALACSYDRELLKDIGKALGEECOSKEVSVILPGCNRKRSPLCORNFFY 132
QY	119 GEDPLLAGAALIRGIOSTGQVATIKHFLCNDQEDRMMVOSIVTERALEYIALPQ 178
Db	133 SEDPYLASMIAISHITKGVQSKGATISLKHFAANNOERHRSVSEIDERLHETYLAAFE 192
QY	179 IAVRDSQGFMAYTINGINVCSCBNPKYDGLRKEWGNDKDRDGTGDSI-HYRTSGDLSLPSVHA 238
Db	193 SVKEAKPWTMCSNKINNLSKKKTLVGPNAQATOVHGGSAALRAY 252
QY	239 GLDLEMPGPFRFRG-ETLKFNNSNGKPFTHIDORAREVLFQVKCAASGVTENGPTTVN 298
Db	253 GLDLEMPGKMKWRSLESVNRNSIQSRSKVTARKVKERAKVVELVKKC-AQGAPETLDGDL 305
QY	299 NTPED-AALRKVGENGIVLKNENNVLSSKKKTLVGPNAQATOVHGGSAALRAY 357

ID	060038	PRELIMINARY;	PRT;	720 AA.	Db	489	-----GRKVITVLLNTGSPVVEVWSRDLVQDGLLYWQA 520
AC	Q50038;						
DT	01-NOV-1996	(TREMBrel. 01, Created)					
DT	01-NOV-1996	(TREMBrel. 01, Last sequence update)					
DE	01-MAR-2002	(TREMBrel. 20, Last annotation update)					
GN	Beta-glucosidase						
OS	Thermotoga neapolitana						
OC	Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.						
OX	NCBI_TaxID=2337;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=ZZ-706-NC24;						
RX	MEDLINE=98048478; PubMed=9387232;						
RA	"Thermotoga neapolitana bgB gene, upstream of lamB, encodes a highly						
RT	thermostable beta-glucosidase that is a laminarinase.";						
RL	Microb. 143:353-354 (1997).						
DR	EMBL: Z77856; CABIN0407_1; -.						
DR	InterPro: IPR001764; GH_3N.						
DR	PFam: PF00933; Glyco_hydro_3_C; 1.						
DR	PRINTS: PR00133; GLYDRASE3.						
DR	PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.						
KW	glycosidase; hydrolase.						
SEQUENCE	720 AA; 81370 MW; 4248A6D2C75A28EF CRC64;						
Query Match	22.4%; Score 982; DB 2; Length 720;						
Best Local Similarity	30.8%; Pred. No. 3.3e-57;						
Matches	265; Conservative 131; Mismatches 275; Indels 190; Gaps 26;						
QY	6 VAIKKLTIAEKVLDLAGIDFW-----HTKALPKHGVPSLRFDPNG 49						
DB	4 VAIILSOTLKEKSERCSG-----GWTSGVWVKSHGWRCKRGETHPVPRVGLPAFLADGPAG 61						
QY	50 VR----GTFKFGNVPACFCFGTSIGSTENOTLLEAEGKMGKAIAKSVAHVLGPTIM 105						
DB	62 LRPNTRENDNTYTTAFAVPEVIMLASTWRLEEVKGAMGEVREYGVDFVLLPAMNI 121						
QY	106 QSPRLGRGFGESIGEDPLLAGAALIRGQIYSTQVATKHFCLNDQEDRMMVQSVT 165						
DB	122 HRPPLCGRNFVEYSEDPVLSGMASSEFVKGQSQGVGACTKHFVANNQETNRMVDTIVI 181						
QY	166 ERALEIYALPFIQIAVDRSOPAFMAYNGINGVSCSENPKYLDGMLRKWGVGDGLMSD 225						
DB	182 ERALEIYALPFIQIAVDRSOPAFMAYNGINGVSCSENPKYLDGMLRKWGVGDGLMSD 241						
QY	226 WGYTISTTEAVAGLDLEMPGP-----RERGETLKENNSNGRFHVIDQARAEVLO 278						
DB	242 WYADNPVHEQKLAGNDLIMPQKAVQVNTTEREEIIMEAALKEGKLSEEVLDCEVRNLK 301						
QY	279 FVKKCAASGTYENTGP-----ETTVNTPET-----HALLRKGNGGIVLKKNNENVLPLSKKK 332						
DB	302 -----VLYNAPSFKNYRSNPKDLEKHAKVAYEAGAEVLLKNE-BALPUSENSK 351						
QY	333 TLIVGPNAKOATYHGGSAALLRAYAVTPFDGLSKQLETPPSYTVGATVPPILGEQL 392						
DB	352 KALFGTQIETIKGGTSQDTHPKYATILEGKTR-----387						
QY	393 TPDGAPGMGRVNNEPGPPTPRQHDLDEFTKDMHLYDYYKARTWADMEGYTAD 452						
DB	388 -----GLN-----FDEELAKIYEDYIKK-----MRETEYKPR-RDSW-----419						
QY	453 EDCTYEGLGVWVCGTAKAYDDLVVNDATRQVCPDAFFGSATREETG-RIMVKGNTK 510						
DB	420 -----GTY-----RPKLSENFLESEKVKL-----AKRNDAVIVYVSSRLSGGYD 460						
QY	511 KFIEGSAAPTILKGOTIVGCHGSILRVGCKKVIDDOAELEKVALHQDQVITCAGLNA 570						
DB	461 RK-----PVKGDPYLS-----DDETDLIKTVS-REFEQ-----488						
QY	571 DWETEGADRASMKLPSVLDQJADVAANPNTVVMOTGIP-EMPWIDATPAV1QAWG 629						
Query Match	21.8%; Score 959; DB 16; Length 721;						
Best Local Similarity	30.4%; Pred. No. 1.3e-55;						
Matches	262; Conservative 130; Mismatches 280; Indels 190; Gaps 26;						
QY	6 VAIKKLTIAEKVLDLAGIDFW-----WHTKALPKHGVPSLRFDPNGY 50						
DB	4 IDELSQSLTTEKVLUWVGLPGLFGNHRVAGAEGETHPVPLGIPAVLADGPAG 63						
QY	51 R----GTFKFGNVPACFCFGTSIGSTENOTLLEAEGKMGKAIAKSVAHVLGPTIM 106						
DB	64 RLPNTRENDNTYTTAFAVPEVIMLASTWRLEEVKGAMGEVREYGVDFVLLPAMNI 123						
QY	107 RSPLGGRGFGESIGEDPLLAGAALIRGQIYSTQVATKHFCLNDQEDRMMVQSVT 166						
DB	124 RHPPLCGRNFVEYSEPVLSGMASSEFVKGQSQGVGACTKHFVANNQETNRMVDTIV 183						
QY	167 RALEIYALPFIQIAVDRSOPAFMAYNGINGVSCSENPKYLDGMLRKWGVGDGLMSD 226						

DR	Pram; PF01915; Glyco_hydro_3_C; 1.
DR	PRINTS; PRO0133; GLYHYDROLASE3.
DR	PROSITE; PS00775; GLYCOSYL_HYDROL_F3; 1.
SQ	SEQUENCE; 828 AA; 86180 MW; 3E8BDB4692F2030D CRC64;
Query Match	21.8%; Score 958.5; DB 2; Length 828;
Best Local Similarity	31.2%; Pred. No. 1.8e-55; Mismatches 355; Indels 93; Gaps. 17;
Matches	262; Conservative 130; Pred. No. 1.8e-55; Mismatches 355; Indels 93; Gaps. 17;
Db	4 IDVEALKLKLTAEKVULLAGIDFWHFKLPRKRGVSLRFDPGNCVRGKFFNGVPPAC 63
Db	54 FPCGTSAGGSTENOTLIEFAGKMGKERAIAKSAHVIWTGPTINMQRSPLGRCFESTEGEDPF 123
Db	64 FPCGTSAGGSTENOTLIEFAGKMGKERAIAKSAHVIWTGPTINMQRSPLGRCFESTEGEDPF 123
Db	67 FPNATLASSANASPLIEKTYAFTTAPEESIGLGEVRLSDGPGVGRGLKFS3RTVAL 66
QY	388 -----RNMKFDDELASLASYEEVKK-----MRETEYKPR-TDSW----- 420
QY	334 LIVGENAKAQTYHGGSAALRAYAATPFDGLSQUELPPPSYTGVAYITVPPPLGEQCLT 393
QY	354 AVFGQIQIEIPIKKGGSGDHTPRYTISTILEGI----- 387
QY	394 PDGAPGMRHVFENNEPPGTNRQHIDELFTKTDMLHVYHPKAADTWADMEGTYADE 453
QY	454 DCTYELGLWVCGTAKAYDDQLVVDNATRQVPGDAFFGSAATREBTG--RINLVGKNTYKF 511
QY	421 -----GTVI-----KPKPENELSEKEKK-----AKKNDVAVVVISRLSISGEGYDR 462
QY	512 KIEFGSAPTYTLKGDTIVPGHSLRVGCKVWIDQAETEKKSWALAK-HDQVITACGLNA 570
QY	463 K-----PVKGDFYL-----DDEELIKTV--SKEFHDQ----- 489
QY	571 DWETEGADRASMSKLPLGCVLDQLIADVAANPNTVWVWQMTGTEB-----PWLDAATPAVQIAWYQ 629
QY	490 -----GKKVWVLNLIGSPLEVASWRLDVGILLWQA 521
Db	124 LAGLAALIRIQCOSTQVQATIHKFLNDQDERRMAMOSVTERALRETAELPQIAVARD 183
Db	127 LTGRLAAVYAVRLQDLYGVAGCILKHLYVANESETERVNMNSTVPAFLRELYKLPPFIAVDE 186
QY	522 GOEMGRIVADLVGKINPSPKLPTPDKYPSMTPVPEPKDNPQRVVTEEDIVGIR 581
QY	687 YVEFADKVNPFPKGHSIYTFFASFLSUSVSKRDKLUSVLSVSKVNGSVRGAQVQOLYKP 745
QY	582 YDFTFGVEPAYFEGYLSYTFKFEYKDLKIDGEPLRVSVTITNGDRACKEVSYVQYK- 640
QY	747 I-QAAKINRNPYELKGFAKE--LQPGETKAVTIEEQKVVAYFDEERDQWVCEGVDYEI 805
QY	641 APKGKIDKPKCFLKAFHKTKLNLPGSEETISLEIPLRDIFASFDKE--WVSEGEYEV 697
QY	806 VSDSSAKDGALRSKFTV-GE 826
QY	698 WGASSR--DIRLDIFLYVEGE 716
RESULT	8
051843	PRELIMINARY; PRT; 828 AA.
051843	051843; 01-JUN-1998 (TREMBrel. 06, Created)
051843	01-JUN-1998 (TREMBrel. 06, Last sequence update)
051843	01-MAR-2002 (TREMBrel. 20, Last annotation update)
051843	Cellobiase.
051843	Cellulomonas biazotea.
051843	Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Micrococcineae; Cellulomonadaceae; Cellulomonas.
051843	SEQUENCE FROM N_A.
051843	STRAIN=ATCC 486;
051843	MEDLINE=9817237; Pubmed=9511746;
051843	EMBL; AR05277; AAC38196; 1.
051843	InterPro; IPR00772; GH_3C.
051843	InterPro; IPR00764; GH_3N.
051843	PFam; PF00933; Glyco_hydro_3; 1.
RESULT	9
087852	PRELIMINARY; PRT; 829 AA.
087852	087852; 01-NOV-1998 (TREMBrel. 08, Created)
087852	01-NOV-1998 (TREMBrel. 08, Last sequence update)
087852	01-JUN-2002 (TREMBrel. 08, Last sequence update)
087852	ERDQVITACGLNAWDWETEGADRASMSKLPLGVLDQLIADVAANPNTVWVWQMTGTPPEMPWL 617
Db	569 QADAVAVVVGVLTEETEESTESVSKTIALPGADALVRAVAAAARRTVWVVAATPVILMPNL 628
QY	618 DATPAVQIAWQGNETGNSIAIDWVFGDYNISGKLSLSPFLRQLDNPAAFLNTRTEAGRTH 677
Db	629 DDDAVLWLPGOBEGHAYAAALGQDPEQPTGRVLYTFRADGAPAW-SVTPVGDLEX 687
QY	678 GEDVYVGYRFYEFADK--DYNPFEGHLSYTFARSNLSVSHKDGKLSLSSVKNNTGSP 735
Db	688 TEGFVGYRGH-WAARRPAPAWFLHGHLGSGYATWEVADATDGDAPATVTVNTGART 746
QY	736 GAQVQVOLYVPLQAQAKINRNPYELKGFAKVELQPGETKAVTIEEQKVVAYFDEERDQW 795
Db	747 SREVVQVYIPLP--ASSDEPVR-LVGVWADATDVAGASARVTVADAR-MWRRWDEAAGGW 801

DE	Putative beta-glucosidase.
SC06597	OR SC8A6_18.
OS	Streptomyces coelicolor.
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetidae; Streptomyctaceae; Streptomyces.
OC	Actinomycetales; Streptomyctaceae; Streptomyces.
NCBI_TaxID	1902;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=A3(2);
RA	Seeber K.J.; Harris D.;
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=A3(2);
RA	Parikhill J., Barrell B.G., Rajandream M.A.;
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=A3(2);
RA	MEDLINE=97000551; PubMed=8843436; Denapaitte D., Eichner A., Cullum J., Kieser H., Hopwood D.A.;
RT	"A set of ordered cosmids and a detailed genetic and physical map for
RT	8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL	Mol. Microbiol. 21:7-36(1996).
RN	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN=A3(2) / M145;
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bakerman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kieser T., Larke L., Murphy L., Oliver T., O'Neil S., Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutten S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;
RT	"Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)." Nature 417:141-147(2002).
RL	EMBL: AL031013; CAA197901; -.
DR	InterPro: IPR002772; GH_3C.
DR	InterPro: IPR001764; GH_3N.
DR	Pfam: PF00933; Glyco_hydro_3_C; 1.
DR	PRINTS: PRO0133; GLYHDRASE3.
SQ	SEQUENCE 859 AA; 89455 MW; BD7EAFEA4FD59F CRC64;
Query Match	21.8%; Score 957; DB 16; Length 859; : Best Local Similarity 29.7%; Pred. No. 2.4e-55; Matches 266; Conservative 131; Mismatches 354; Indels 146; Gaps 22;
RESULT 10	070021
ID	070021
AC	PRELIMINARY;
DT	01-AUG-1998 (TREMBLrel. 07, Created)
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT	01-MAR-2005 (TREMBLrel. 20, Last annotation update)
DE	Beta-glucosidase.
GN	ERYB1.
OS	Saccharopolyspora erythraea (Streptomyces erythraeus).
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Streptomyctaceae; Streptomyces.
OC	Actinomycetales; Pseudonocardinae; Pseudonocardiaceae; Saccharopolyspora.
OC	Saccharopolyspora.
NCBI_TaxID	1836;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=NRRL2338;
RA	MEDLINE=98273631; PubMed=9613575;
RA	Gaisser S., Boehm G.F., Doumith M., Raynal M.C., Dhillon N., Corriss J., Leadlay P.F.;
RA	"Analysis of EryB1 and EryB2 from the erythromycin biosynthetic gene cluster in Saccharopolyspora erythraea." Mol. Gen. Genet. 258:78-88(1998).
RL	EMBL: Y1-327; CAA47022; -.
DR	InterPro: IPR002772; GH_3C.
DR	InterPro: IPR001764; GH_3N.
DR	Pfam: PF00933; Glyco_hydro_3_C; 1.
DR	PRINTS: PRO0133; GLYHDRASE3.
SQ	SEQUENCE 808 AA; 86677 MW; 0C619016CC00A751 CRC64;
Query Match	21.7%; Score 954; DB 2; Length 808; Best Local Similarity 30.8%; Pred. No. 3.4e-55;

Matches	260;	conservative	128;	Mismatches	334;	Indels	122;	Gaps	21;	OC
QY	6	VEAILKKITLAKVVDLILAGIDFWH-----		TKALPKHGVPSLRFDGPNGVRGT	54					OC
Db	37	WELIGOLTLDEKLISFV-----	WWDYNEKDPLAKLWLPGVFRGIPQIGTDPAGVTHQ	92						Caulobacter
QY	55	FFNGVPAFCFGTSLGSTENOTLLEAGKAMKGKEIAKSAHNVILGPTTINMQRSLGGRC	114							alpha subdivision: Caulobacter group;
Db	93	-----PAIAMPAPVALASADEFDRALAHYGTWLGREGRAFFODDILGPWMVNRYVQAGRN	147							NCBI_TAXID=155892;
QY	115	FESIGEDPFLAGGAAALIRGICSTGQVATKHFCLNDQEDDRMAMQSVITERAIREYA	174							[1]
Db	148	FETFSEDEPLVARTAAQIRGIDHSQGLMTSAKHYANTQITDFTIDVVDQDRLRELE	207							RC
QY	175	LFQIAVDRDQSQFGMFTAYNGTNGVSCSENPKYDQMLRKEWGMGLIMSDWYCSYSTE	234							RE
QY	208	PCFEAAVA-AGATSVMACAYPKVNGTHACGHROLILBILKEQWGFKGWVMSDWYATHA-TE	265							RE
QY	235	AVVAGLDEMLEMPPR-----	-----	-----	-----	-----	-----			RE
Db	266	DLVAGLQDQEMGVEVREDGSFLRKGKLYGEALKAIAREGRIPESALDAVRRLTDFERGL	325							RE
QY	286	SGVTTENGPEPTVWNNTETAAALLRKVNGEGLVLLKNNENVPL- SKKKTLLIVPNAKA	343							RE
Db	326	--LDETTRPRPERDVAGGTRIAQEAEASGAVLRLNEGGVPLPDRPAGQIAVIGPSAQD	383							RE
QY	344	TYHGGSAALRAYAYATVPDFGSLKOLET - PPSYVGAATVPPILGEQCLTPDGAPGMR	401							RE
Db	384	KVTGLGSSYVRFANAPLDTITQPGSGGRVGYSGVEELKGAPI-PETALQPAFVG--	440							RE
QY	402	WVFNNEPPTGTRNQHIDELFTKTDMLVYHPKRADTWADMEGTYADECTYEGL	461							RE
Db	441	-----EVIP-----								RE
QY	462	WVCGTAKAYVDDQLVVDNATKQVPSDAFFGCSATREBETGRINLVKGNTYFKIPEG--	518							RE
Db	469	RIDG-----								RE
QY	519	PTYTLKGDTIVPGHGLSLRYVGCKVYD-----	-----	-----	-----	-----	-----			RE
Db	498	PWLTKGHEHTIQMTGAPVGGSLDQVDTWVTPGHAQREFDAVERARDSDVAVFA--	554							RE
QY	570	ADWETEGADRSKMKPGVLDQILIAWAAAPNTIVVMQGTPPEMPWIDATPAVIOAWG	629							RE
Db	555	YDGAATTADETRSLSLPGTQOKLIDAWAANPNTIVVMQGTPPEMPWIDATPAVIOAWG	614							RE
QY	630	GNETGNSIADVFGDYNPSKSLISFPKRLQDNIF-----	-----	-----	-----	-----	-----			RE
Db	615	GOAGAEATTALFGAEPGRGLTQTFPSQERTVFGDGFARFQVGDGVHISGIFSGR	674							RE
QY	687	YEEFAKDVWPFHGLSYTAFSNSLPSYHCKOCLSLSVKNTGSGPQAOQLYVKP	746							RE
Db	675	WIDREGVDPFLPFHGLSYTTERDIDVWVTRDGLDVWVTRVNGQRRGSDVWQVYIGP	734							RE
QY	747	LOQAKINRNPYKELKEFEKLEQETKAVTIEQEKVYAYFEERDOWCVERGDEVIV	806							RE
Db	735	SPQVPLDQAPRQLAGYOKVELAPGETKRVYHWAER-ALQHWDIAAGGWLKGKRAVEI	793							RE
QY	807	SDDS 810								RE
Db	794	GSSS 797								RE
RESULT	11									
Q9A7G7		PRELIMINARY;	PRT:	762 AA.						
ID	Q9A7G7									
AC	Q9A7G7;									
DT	01-JUN-2001 (TREMBrel. 17, Created)									
DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)									
DE	Beta-D-glucosidase.									
GN	CC1756.									
OS	Caulobacter crescentus.									
Query	Match	21.6%	Score	946.5;	DB	16;	Length	762;		OC
Best	Local Similarity	30.5%	Score	946.5;	DB	16;	Length	762;		OC
Matches	257;	Conservative	115;	Mismatches	273;	Indels	199;	Gaps	20;	OC
QY	4	IDVEAILKKLTAEKVVDLAGI-----	-----	-----	-----	-----	-----			OC
Db	55	VRADALVKQMTLDEKITYLHGLEPPMAKPPAPIDMDMISAGVYVPGVPLNINPIRESASLG	114							OC
QY	50	VRG-TKEFGVPAFCFGTSLGSTENOTLLEAGKAMKGKEIAKSAHNVILGPTTINMQRSLGGRC	108							OC
Db	115	VANQVEQRKGDVATALPSGLALASTIFPKLADFGGAMIGEARAKTIVFNLVLAGGYNLTRD	174							OC
QY	109	PIGGRGEFSIGEDPFLAGGAAALIRGICSTGQVATKHFCLNDQEDDRMAMQSVITERA	168							OC
Db	175	PWAGRNFETFLGEDPLLAGEMVAEQIKVQSNRIVSTKHFALNAOTBGR;YMDQIDEAD	234							OC
QY	169	LREIYALPFQIAVDRSDPSGAFMFTAYNGTNGVSCSENPKYDQMLRKEWGMGLIMSDWY	228							OC
Db	235	RESDLIAFQIAKEISNPASVNCAYVNKGWACENDFLNPLSKK-KRTLIVGPNAK	294							OC
QY	229	TSTTEAVAVAGLDEMLEMPPR-----	-----	-----	-----	-----	-----			OC
Db	295	VHSTVKALLAGD-QQSGLDQTIFGDEKIAVAKVGEKSQARVDDMVRKILHGV-1	350							OC
QY	285	ASCVTENGPEPTVNN-NTPEPALLRKVNGEGLVLLKNNENVPLSKK-KRTLIVGPNAK	341							OC
Db	351	SSGLMDQIPTPSAQDIDYDAHKAQVTAVERGVSILKNDRGLPLSKRASKRIVLGAHD	410							OC
QY	342	QATYHGGSAALRAYAYATVPDFGSLKOLETIPPSYVGAATVPPILGEQCLTPDGAPGMR	401							OC
Db	411	VGVISGGSSQV-----								OC
QY	402	WVFNNEPPTGTRNQHIDELFTKTDMLVYHPKRADTWADMEGTYADECTYEGL	461							OC
Db	442	-----								OC
QY	462	WVCGTAKAYVDDQLVVDNATKQVPSDAFFGCSATREBETGRINLVKGNTYFKIEGSAPTY	521							OC
Db	450	-----								OC
QY	522	TUKGDTIVPGHGLSLRYVGCKVYDDEIEKSVLAKVNGEGLVLLKNNENVPLSKK-KRTLIVGPNAK	581							OC
Db	455	AIKA-----ANPSEA-----	502							OC
QY	582	MKLPGVLDQIADVAAANPNTVWVWVQGTPPEMPWIDATPAVIOAWGIGNETGNSIADV	641							OC

Db	503	IALPENODALIERAUSAANKNAVWVILETGPVILPWNLDKVGAVLAWYFGORGQSAIRL	562	QY	283	CASGYTENGETTVNTPTAAFLRKVGNEGTVWILKNNNNVPLS- KKKYLVIPNA	340
Db	642	FGYINPSSKLISLSPKQLQDNP-----AFINERTAGR-----TLYGEDVY	682	QY	344	GULLATPAPRE-- RDKAGQAVSRKAENAVILRNQGQALPLAGDKASTAVIGPTA	400
Db	563	FGYINPSSKLISLSPKQLQDNP-----AFINERTAGR-----TLYGEDVY	682	QY	341	KQATYHGGSAHLARAYAVVPPDGUSQLETPPSVTVGAYTVPILBQCTPDGAGM	400
Db	683	VGRRYEFADKDVFNPFGHGSYTTAFSNLSVSHKGKLSVLSVKNITGSVPAQAGL	742	QY	401	VDPKVIGLGSAHWVPPSAASALDTIKAR-----AGAGATVYETGETFGTOIPAGN	452
Db	623	VGTRWFAEQKRRRLYPPGQYLSUTSGYKNKVEDDG-LKVSEDTNTGKVQAGATPQL	681	QY	401	RWRFVNBPPTPNRQHIDELFFTKTDMHLYDYYHKAADTWADMEGYTADDECTYLG	460
QY	743	YVKPLQAAKINRPRVKEKGAKYELQGETKAVTIEQKVAATFDEEDQWYERGY	802	QY	453	LSPAFNQG-----HOLE--PGKAGALY--DGTLVTPADGEYRIL	487
Db	682	YVSGO---RKAMRLLAGFQKVLARGEKTVNLNEPRILADY-DTAKPGWITAGTY	736	QY	803	EVIV 806	
Db	737	PLVY 740		QY	737	PLVY 740	
RESULT	12			QY	461	LWVCGTAKAYVDDQQLVVDNATKOPDADFAFRSARATBETGRNLUVGNTYKFKIEGSA	520
ID	Q9ZGH5	PRELIMINARY;	PRT;	QY	488	VRATG--GYATVOL---GSHTIEAQVQVKS--SPLIKLIKG-THRL-----	527
AC	Q9ZGH5;	01-MAY-1999 (TREMBIrel. 10, Created)		Db	521	YTLKGDTIVPGHSLRVGGCKVDDAEIESVLAKEHODVITCAGLNADWETEGADRA	580
DT	01-MAY-1999	(TREMBIrel. 10, Last sequence update)		QY	528	-TISGFAMASATPLSLELGWVTPAADDATIKAVERAKARTAVFA--YDGTGEGVDRP	583
DT	01-MAR-2002	(TREMBIrel. 20, Last annotation update)		QY	581	SMKLPGVILDQDADVAANPNTVVMQGTPBEMPLDAPTAIVQAWYSGNETENSLAD	640
DE	Beta-glucosidase.			Db	584	NLSLPGTQDKLISAVADANPNTVWLTGQVSLMPKSLTRAVLDWYPGZAGAATAAL	643
GN	DESR.			QY	641	VFGYINPSSKLISLSPKQLQDNPFLAFLNTERG-----TLYGEDVY	590
OS	Streptomyces venezuelae.			Db	644	LYGDVNPSGKLTQSF-----PAENOHAVAGDPTSYPGVQDNOQYQYREGIHGYRWFED	696
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			QY	691	ADKVNFPVPGHGSYTTAFSNLSVSH-KDGLKUSVISUSVKNTGSVPAQAOQLYKKPLQ	749
OX	NCBI_TAXID=54571;			Db	697	ENVKPLFPFGHGSYTSFQSAFTVWVTRISTGGKLVTVNRSKRGQEVQYAOYIGASP	756
RP	SEQUENCE FROM N.A.			QY	750	AKINRPKLKGAKYELQGETKAVTIEQKVAATFDEEDQWYERGY	809
RC	STRAIN=ATCC15439;			Db	757	VTAPOAKKLVGYTVKVLAGEAKTVTVN-----DRQIQTGS	795
RK	MIDDLE=984533;			QY	810	SAAKGVALGVKFTIV 824	
RA	Xue Y., Zhao L., Liu H.W., Sherman D.H.;			Db	796	SSAD---LNGSATV 806	
RT	"A gene cluster for macrolide antibiotic biosynthesis in streptomycetes venezuelae: architecture of metabolic diversity".			RESULT	13		
RL	Proc. Natl. Acad. Sci. U.S.A. 95:12111-12116(1998).			ID	Q9F3Y0	PRELIMINARY;	PRT;
DR	EMBL: AE079762; AACB679.1;			AC	Q9F3Y0;		
DR	InterPro; IPR003772; GH_3C.			DT	01-MAR-2001	(TREMBIrel. 15, created)	
DR	InterPro; IPR001764; GH_3N.			DT	01-MAR-2001	(TREMBIrel. 16, Last sequence update)	
DR	Pfam; PF01915; Glyco_hydro_3_C; 1.			DT	01-MAR-2002	(TREMBIrel. 20, Last annotation update)	
DR	PRINTS; PRO01133; GLYHDRASE3.			DE	EXO-1,4-beta-glucosidase (EC 3.2.1.21).		
DR	SEQUENCE FROM N.A.			DR	EXO-1,4-beta-glucosidase (EC 3.2.1.21).		
DR	STRAIN=ATCC15439;			DR	Prevotella albensis.		
DR	Middle=984533;			DR	Bacteroides; Bacteroidetes; Bacteroidales; Prevotellaceae;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Prevotella; Prevotella albensis.		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	NCBI_TAXID=77768;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	SEQUENCE FROM N.A.		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	STRAIN=M384;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Best Local Similarity 32.2%; Pred. No. 1.5e-54;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;			DR	Walker N.D., McEwan N.R., Wallace R.J.;		
DR	Score 21.5%; Score 944.5; DB 2; Length 809;			DR	Walker N.D., McEwan N.R., Wallace		

521 AGNAVADILIGKVNPGSKUPLTFKRLIEDSPSPGNPNAKTIYGEIYGYRFDTKR 580
 QY 693 KDVNPFPFGHGLSYTTFAFSNLVS-----HKQKLSLSSLVNTGSPGQAOLYKP 746
 Db 581 IAPMPFPFGLSSYTVNGKISPERKPVNIDTENSIEVSIPVRNTSGIDTKEVQVYHD 640
 QY 747 LQAAKINRPUKELGFAKVELQFETKAVTIEOEKYVAYFDEERDOWCVERGDYEVIV 806
 Db 641 -NAAEQLDREQEQLAKEVKVSLAGEEKVAF-KLDRKAFSYVEDKNWWLVEGLFTL 698
 QY 807 SDSS 810
 Db 699 GRSS 702

RESULT 15
 Q9RH03 PRELIMINARY; PRT: 732 AA.
 AC Q9RH03;
 DT 01-MAY-2000 (TREMBirel. 13, Created)
 DT 01-MAR-2002 (TREMBirel. 20, Last annotation update)
 DE Beta glucosidase precursor.
 RN 1
 OS Azospirillum irakense.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
 OC Azospirillum;
 NCBI_TAXID=34011;
 RN 2
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC1;
 RA Faure D., Desair J., Keijers V., Proost P., Henrissat B.,
 RA Vanderleyden J.;
 RT "The sala" and salB genes encode two beta-glucosidases in Azospirillum
 RT irakense.";
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF09429; AAC21798 1; -.
 DR InterPro; IPR0272; GI_3C.
 DR InterPro; IPR001764; GI_3C.
 DR Pfam; PF00933; Glyco_hydro_3 1.
 DR Pfam; PF01915; Glyco_hydro_3_C 1.
 DR PRINTS; PR00133; GHYDRASE3.
 KW Signal.
 FT SIGNAL 1 26 POTENTIAL
 SQ SEQUENCE 732 AA; 77650 MW; F792A947923F4E1E CRC64;

Query Match 20.9%; Score 919.5; DB 2; Length 732;
 Best Local Similarity 30.7%; Pred. No. 6e-53; Mismatches 260; Conservative 106; Indels 205; Gaps 23;
 Matches 49 ELTLIDERISLHGPMLPFPGSPIPEPSLWVIF--PGVRLGIPALKETDASLCV- 104

QY 12 KLTIAEKVDSLQG-----IDFWHITKALFRKHGVPSLRFEDGPNGVR 51
 Db 105 -TNPKNVRPGDTATALPSGLALASTENPKLSYDGGAAKEAMASKGFNVLLAGGANLARD 163
 QY 109 PLGGRGFESIGEDEFLAGIGAALIRGOSTGQATIKHFLNDQEDRRMAMYQSVIYERA 168
 Db 164 PRGRNRFYELGEDPFLLAGLIESIRGQTSQNTISTVHFSLSINGQETNRHMGNSVIDAA 223
 QY 169 LREIYALPQIAVRDQSOPGAFMAYNGNGVSESENPKVLDGMRLKEWGWGDLIMSIDWYG 228
 Db 224 HRESQDQIAQIATERGQPGCYSMCAVNLNGAYSCGNDHILNKVLKGDWGKGVMSB-W-G 282
 QY 229 TYSTPEAVAGLDEIEMPGPR---FRGTTKVNNSQKPFHIDORAREVLFVRC 284
 Db 283 AVPATDFALKGLD-QOSGGOLDEKIWFGLDLIKEAAGTIPAPERLSMSRRL--RSMF 338
 QY 285 ASGYIENGEBETTIVNTPEPAALLKVGNGIVLKLNNENNLPLSK-KKTLWGPNAKQA 343
 Db 339 AAGFFDGKPGKPVVLDAAHAATAKOVADEGTIVLILLANDKGLLPLAAGSOKIAVIGGADQ 398

QY 344 THHGGGAAALRAYVAVTPFDGLSKOLETPSPYTVGAYTTVPIPLGEOCLPDGADGMRWR 403
 Db 399 VLSGAGSSQVT-----SVGGNPVWIVPGGGLM--AAFLRQA 433
 QY 404 VNEEPCTPQRQHIDELFFTKTDMHLDYVYHPKAAUTWYALMEGTYTADPCTYELGLVY 463
 Db 434 YHNSSPLKALKERL-----PNATIRF---NDGRYSA----- 461
 QY 464 CGTAKAVVDQLVWDNATKQPGDAFFGSAATRETEGRINLVKGNTYKFERGSAPTYL 523
 Db 462 ----- 461

QY 524 KGDITIVPGHGLSLRVGGCKVDDQAEIETKSVALKERHDQVILICAGLNADWETEGADRASM 583
 Db 462 -----AAALARQSDIVILFAN--QWMSEGMDAYDL 491

QY 584 LPGVLDLADVAANPNTVYVMOTGTPPEIPWLDATPAV1QAWYGGNEEGNSIADWVG 643
 Db 492 LPQGODALITEA VAEANPNAVIVLQTGGPVLMPWKDKWVGA TWSAWGSKQGEAIDLVG 551
 QY 644 DYNPSGKLISLSPKRLQ---DNPRLNFRTEAGRLYGEDVYGYRYEFADKVN- 696
 Db 552 KINPSSGRLSPTRPASQDQPHFEPWNLRIKQFQDWFYEGSDIGYR-RFAAKGMKU 609
 QY 697 FPPFGHGLSYTTFAFNSLUSVSHKDQ-LSVSISVKNIGSPGQAOLYVPLQAAKINRP 755
 Db 610 FPPFGHGLSYTTFAYDKLVY--KGGERLEVSQVTINGKLGQDKAQIYLAGANGOKLQR- 666
 QY 756 VELKGAKVELQFETKAVTIEQEKYVAYFDEERDOWCVERGDYEVIVSDSSAAKDG 815
 Db 667 ---LIGFEKIDLKPGERTTIVIKADPRLAR-FDEQGHOWRIDDGDYDVWGRSATM-- 719
 QY 816 VALRGK 821
 Db 720 TVLSGK 725

Search completed: April 26, 2003, 13:09:46
 Job time : 52 secs

